SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS:
    (A) NAME: Peter Stougaard
            Ole Cai Hansen

(ii) TITLE OF INVENTION: RECOMBINANT HEXOSE OXIDASE, A
         METHOD OF PRODUCING SAME AND USE OF SUCH ENZYME

(iii) NUMBER OF SEQUENCES: 34

(iv) CORRESPONDENCE ADDRESS:
    (A) ADDRESSEE: Hunton & Williams
    (B) STREET: 1900 K Street, N.W.
    (C) CITY: Washington
    (D) STATE: D.C.
    (E) COUNTRY: U.S.A.
    (F) ZIP: 20006-1109

(v) COMPUTER READABLE FORM:
    (A) MEDIUM TYPE: Floppy disk
    (B) COMPUTER: IBM PC compatible
    (C) OPERATING SYSTEM: PC-DOS/MS-DOS
    (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
    (A) APPLICATION NUMBER: 08/669,304
    (B) FILING DATE: 12 July 1996
    (C) CLASSIFICATION:

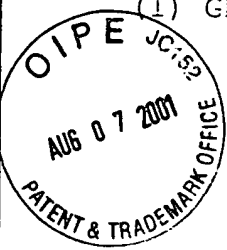(vii) PRIOR APPLICATION DATA: -
    (A) APPLICATION NUMBER: 08/476,910
    (B) FILING DATE: 7

(viii) ATTORNEY/AGENT INFORMATION:
    (A) NAME: Stanislaus Aksman
    (B) REGISTRATION NUMBER: 28,562
    (C) REFERENCE/DOCKET NO.: 54320.000003

(ix) TELECOMMUNICATION INFORMATION:
    (A) TELEPHONE: (202) 955-1926
    (B) TELEFAX: (202) 778-2201
    (C) TELEX: None

(1) INFORMATION FOR SEQ ID NO: 1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 8 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Tyr Glu Pro Tyr Gly Gly Val Pro
1   5

(2) INFORMATION FOR SEQ ID NO: 2:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Ala Ile Ile Asn Val Thr Gly Leu Val Glu Ser Gly Tyr Asp Xaa Xaa
1   5   10   15

Xaa Gly Tyr Xaa Val Ser Ser
         20

(2) INFORMATION FOR SEQ ID NO: 3:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

   (ii) MOLECULE TYPE: peptide

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Asp Leu Pro Met Ser Pro Arg Gly Val Ile Ala Ser Asn Leu Xaa Phe
1   5   10   15

(2) INFORMATION FOR SEQ ID NO: 4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 14 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Asp Ser Glu Gly Asn Asp Gly Glu Leu Phe Xaa Ala His Thr
1   5   10

(2) INFORMATION FOR SEQ ID NO: 5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Tyr Tyr Phe Lys
1

(2) INFORMATION FOR SEQ ID NO: 6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 15 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Asp Pro Gly Tyr Ile Val Ile Asp Val Asn Ala Gly Thr Xaa Asp
1   5   10   15

(2) INFORMATION FOR SEQ ID NO: 7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 11 amino acids
        (B) TYPE: amino acid

```
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

    Leu Gln Tyr Gln Thr Tyr Trp Gln Glu Glu Asp
    1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 8:

```
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

    Xaa Ile Arg Asp Phe Tyr Glu Glu Met
    1               5
```

(2) INFORMATION FOR SEQ ID NO: 9:

```
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide    -

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

    Ala Ile Ile Asn Val Thr Gly Leu Val Glu Ser Gly Tyr Asp Xaa Xaa
    1               5                   10                  15

    Xaa Gly Tyr Xaa Val Ser Ser
                20
```

(2) INFORMATION FOR SEQ ID NO: 10:

```
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown
```

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```
Asp Leu Pro Met Ser Pro Arg Gly Val Ile Ala Ser Asn Leu Trp Phe
1               5               10              15
```

2) INFORMATION FOR SEQ ID NO: 11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 14 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

```
Asp Ser Glu Gly Asn Asp Gly Glu Leu Phe Xaa Ala His Thr
1               5               10
```

(2) INFORMATION FOR SEQ ID NO: 12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

```
Tyr Tyr Phe Lys
1
```

(2) INFORMATION FOR SEQ ID NO: 13:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 15 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: unknown
        (D) TOPOLOGY: unknown

    (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

Asp Pro Gly Tyr Ile Val Ile Asp Val Asn Ala Gly Thr Pro Asp
1   5   10   15

) INFORMATION FOR SEQ ID NO: 14:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 11 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: unknown
       (D) TOPOLOGY: unknown

  (ii) MOLECULE TYPE: peptide

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

   Leu Gln Tyr Gln Thr Tyr Trp Gln Glu Glu Asp
   1   5   10

(2) INFORMATION FOR SEQ ID NO: 15:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 9 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS: unknown
       (D) TOPOLOGY: unknown

  (ii) MOLECULE TYPE: peptide

  (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

   Xaa Ile Arg Asp Phe Tyr Glu Glu Met
   1   5   -

(2) INFORMATION FOR SEQ ID NO: 16:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 20 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

  (ii) MOLECULE TYPE: other nucleic acid

  (ix) FEATURE:
       (A) NAME/KEY: modified base; N=inosine
       (B) LOCATION: base pairs 3, 6 and 12
       (C) IDENTIFICATION METHOD: commercially available
       (D) OTHER INFORMATION

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

YTNGTNGARW SNGGNTAYGA   20

(2) INFORMATION FOR SEQ ID NO: 17:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 23 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (ix) FEATURE:
       (A) NAME/KEY: modified base; N=inosine
       (B) LOCATION: base pairs 6 and 12
       (C) IDENTIFICATION METHOD: commercially available
       (D) OTHER INFORMATION

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

AACCANARRT TNGANGCDAT NAC   23

(2) INFORMATION FOR SEQ ID NO: 18:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 23 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (ix) FEATURE:
       (A) NAME/KEY: modified base; N=inosine
       (B) LOCATION: base pairs 6 and 15
       (C) IDENTIFICATION METHOD: commercially available
       (D) OTHER INFORMATION

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GARGGNAAYG AYGGNGARCT NTT   23

(2) INFORMATION FOR SEQ ID NO: 19:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 23 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single

          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: other nucleic acid

     (ix) FEATURE:
          (A) NAME/KEY: modified base; N=inosine
          (B) LOCATION: base pairs 3 and 9
          (C) IDENTIFICATION METHOD: commercially available
          (D) OTHER INFORMATION

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

AANAGYTCNC CRTCRTTNCC YTC   23

(2) INFORMATION FOR SEQ ID NO: 20:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 22 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: other nucleic acid

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:
ATTGGGGCTC CTTCAAGACC TT   22

(2) INFORMATION FOR SEQ ID NO: 21:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 18 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear   -

     (ii) MOLECULE TYPE: other nucleic acid

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

TGATGATTCC AAAGTTTC   18

(2) INFORMATION FOR SEQ ID NO: 22:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 18 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

TTGGAAGAAT ACGGTTGG   18

(2) INFORMATION FOR SEQ ID NO: 23:

 (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 20 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

 (ii) MOLECULE TYPE: other nucleic acid

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

TACTATTTCG TCTGCTTGGG   20

(2) INFORMATION FOR SEQ ID NO: 24:

 (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 20 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

 (ii) MOLECULE TYPE: other nucleic acid

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GAACTCTTCC GTGGTCTCCT   20

(2) INFORMATION FOR SEQ ID NO: 25:

 (i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 20 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

 (ii) MOLECULE TYPE: other nucleic acid

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

CCACCTGCGT GTTGGGGTCT   20

(2) INFORMATION FOR SEQ ID NO: 26:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

CAGATCTACA AAACATGCGA G   21

(2) INFORMATION FOR SEQ ID NO: 27:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 18 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

TGTCGCAGAC TGTACTTG   18

(2) INFORMATION FOR SEQ ID NO: 28:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 18 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single -
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GAGTGTACAC GACATAAA   18

(2) INFORMATION FOR SEQ ID NO: 29:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

ATGGCTACTC TTCCCCAGAA AG   22

(2) INFORMATION FOR SEQ ID NO: 30:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1801 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION:84..1721

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

TGAATTCGTG GGTCGAAGAG CCCTTTGCCT CGTCTCTCTG GTACCGTGTA TGTCAAAGGT   60

TCGCTTGCAC ACTGAACTTC ACG ATG GCT ACT CTT CCT CAG AAA GAC CCC   110
                         Met Ala Thr Leu Pro Gln Lys Asp Pro
                         1   5

GGT TAT ATT GTA ATT GAT GTC AAC GCG GGC ACC GCG GAC AAG CCG GAC   158
Gly Tyr Ile Val Ile Asp Val Asn Ala Gly Thr Ala Asp Lys Pro Asp
 10   15   20   25

CCA CGT CTC CCC TCC ATG AAG CAG GGC TTC AAC CGC CGC TGG ATT GGA   206
Pro Arg Leu Pro Ser Met Lys Gln Gly Phe Asn Arg Arg Trp Ile Gly
              30   35   40

ACT AAT ATC GAT TTC GTT TAT GTC GTG TAC ACT CCT CAA GGT GCT TGT   254
Thr Asn Ile Asp Phe Val Tyr Val Val Tyr Thr Pro Gln Gly Ala Cys
          45   50   55

ACT GCA CTT GAC CGT GCT ATG GAA AAG TGT TCT CCC GGT ACA GTC AGG   302
Thr Ala Leu Asp Arg Ala Met Glu Lys Cys Ser Pro Gly Thr Val Arg
       60   65   70

ATC GTC TCT GGC GGC CAT TGC TAC GAG GAC TTC GTA TTT GAC GAA TGC   350
Ile Val Ser Gly Gly His Cys Tyr Glu Asp Phe Val Phe Asp Glu Cys
      75   80   85

```
GTC AAG GCC ATC ATC AAC GTC ACT GGT CTC GTT GAG AGT GGT TAT GAC        398
Val Lys Ala Ile Ile Asn Val Thr Gly Leu Val Glu Ser Gly Tyr Asp
 90                  95                  100                 105

GAC GAT AGG GGT TAC TTC GTC AGC AGT GGA GAT ACA AAT TGG GGC TCC        446
Asp Asp Arg Gly Tyr Phe Val Ser Ser Gly Asp Thr Asn Trp Gly Ser
                     110                 115                 120

TTC AAG ACC TTG TTC AGA GAC CAC GGA AGA GTT CTT CCC GGG GGT TCC        494
Phe Lys Thr Leu Phe Arg Asp His Gly Arg Val Leu Pro Gly Gly Ser
                 125                 130                 135

TGC TAC TCC GTC GGC CTC GGT GGC CAC ATT GTC GGC GGA GGT GAC GGC        542
Cys Tyr Ser Val Gly Leu Gly Gly His Ile Val Gly Gly Gly Asp Gly
                 140                 145                 150

ATT TTG GCC CGC TTG CAT GGC CTC CCC GTC GAT TGG CTC AGC GGC GTG        590
Ile Leu Ala Arg Leu His Gly Leu Pro Val Asp Trp Leu Ser Gly Val
                 155                 160                 165

GAG GTC GTC GTT AAG CCA GTC CTC ACC GAA GAC TCG GTA CTC AAG TAT        638
Glu Val Val Val Lys Pro Val Leu Thr Glu Asp Ser Val Leu Lys Tyr
170                 175                 180                 185      -

GTG CAC AAA GAT TCC GAA GGC AAC GAC GGG GAG CTC TTT TGG GCA CAC        686
Val His Lys Asp Ser Glu Gly Asn Asp Gly Glu Leu Phe Trp Ala His
                 190                 195                 200

ACA GGT GGC GGT GGC GGA AAC TTT GGA ATC ATC ACC AAA TAC TAC TTC        734
Thr Gly Gly Gly Gly Gly Asn Phe Gly Ile Ile Thr Lys Tyr Tyr Phe
                 205                 210                 215

AAG GAT TTG CCC ATG TCT CCA CGG GGC GTC ATC GCA TCA AAT TTA CAC        782
Lys Asp Leu Pro Met Ser Pro Arg Gly Val Ile Ala Ser Asn Leu His
                 220                 225                 230

TTC AGC TGG GAC GGT TTC ACG AGA GAT GCC TTG CAG GAT TTG TTG ACA        830
Phe Ser Trp Asp Gly Phe Thr Arg Asp Ala Leu Gln Asp Leu Leu Thr
                 235                 240                 245

AAG TAC TTC AAA CTT GCC AGA TGT GAT TGG AAG AAT ACG GTT GGC AAG        878
Lys Tyr Phe Lys Leu Ala Arg Cys Asp Trp Lys Asn Thr Val Gly Lys
250                 255                 260                 265

TTT CAA ATC TTC CAT CAG GCA GCG GAA GAG TTT GTC ATG TAC TTG TAT        926
Phe Gln Ile Phe His Gln Ala Ala Glu Glu Phe Val Met Tyr Leu Tyr
                 270                 275                 280
```

```
ACA TCC TAC TCG AAC GAC GCC GAG CGC GAA GTT GCC CAA GAC CGT CAC    974
Thr Ser Tyr Ser Asn Asp Ala Glu Arg Glu Val Ala Gln Asp Arg His
            285             290             295

TAT CAT TTG GAG GCT GAC ATA GAA CAG ATC TAC AAA ACA TGC GAG CCC   1022
Tyr His Leu Glu Ala Asp Ile Glu Gln Ile Tyr Lys Thr Cys Glu Pro
            300             305             310

ACC AAA GCG CTT GGC GGG CAT GCT GGG TGG GCG CCG TTC CCC GTG CGG   1070
Thr Lys Ala Leu Gly Gly His Ala Gly Trp Ala Pro Phe Pro Val Arg
    315                 320             325

CCG CGC AAG AGG CAC ACA TCC AAG ACG TCG TAT ATG CAT GAC GAG ACG   1118
Pro Arg Lys Arg His Thr Ser Lys Thr Ser Tyr Met His Asp Glu Thr
330                 335             340             345

ATG GAC TAC CCC TTC TAC GCG CTC ACT GAG ACG ATC AAC GGC TCC GGG   1166
Met Asp Tyr Pro Phe Tyr Ala Leu Thr Glu Thr Ile Asn Gly Ser Gly
            350             355             360

CCG AAT CAG CGC GGC AAG TAC AAG TCT GCG TAC ATG ATC AAG GAT TTC   1214
Pro Asn Gln Arg Gly Lys Tyr Lys Ser Ala Tyr Met Ile Lys Asp Phe
            365             370             375

CCG GAT TTC CAG ATC GAC GTG ATC TGG AAA TAC CTT ACG GAG GTC CCG   1262
Pro Asp Phe Gln Ile Asp Val Ile Trp Lys Tyr Leu Thr Glu Val Pro
            380             385             390

GAC GGC TTG ACT AGT GCC GAA ATG AAG GAT GCC TTA CTC CAG GTG GAC   1310
Asp Gly Leu Thr Ser Ala Glu Met Lys Asp Ala Leu Leu Gln Val Asp
    395                 400             405

ATG TTT GGT GGT GAG ATT CAC AAG GTG GTC TGG GAT GCG ACG GCA GTC   1358
Met Phe Gly Gly Glu Ile His Lys Val Val Trp Asp Ala Thr Ala Val
410                 415             420             425

GCG CAG CGC GAG TAC ATC ATC AAA CTG CAG TAC CAG ACA TAC TGG CAG   1406
Ala Gln Arg Glu Tyr Ile Ile Lys Leu Gln Tyr Gln Thr Tyr Trp Gln
            430             435             440

GAA GAA GAC AAG GAT GCA GTG AAC CTC AAG TGG ATT AGA GAC TTT TAC   1454
Glu Glu Asp Lys Asp Ala Val Asn Leu Lys Trp Ile Arg Asp Phe Tyr
            445             450             455

GAG GAG ATG TAT GAG CCG TAT GGC GGG GTT CCA GAC CCC AAC ACG CAG   1502
Glu Glu Met Tyr Glu Pro Tyr Gly Gly Val Pro Asp Pro Asn Thr Gln
            460             465             470
```

```
GTG GAG AGT GGT AAA GGT GTG TTT GAG GGA TGC TAC TTC AAC TAC CCG   1550
Val Glu Ser Gly Lys Gly Val Phe Glu Gly Cys Tyr Phe Asn Tyr Pro
    475             480             485

GAT GTG GAC TTG AAC AAC TGG AAG AAC GGC AAG TAT GGT GCC CTC GAA   1598
Asp Val Asp Leu Asn Asn Trp Lys Asn Gly Lys Tyr Gly Ala Leu Glu
490             495             500             505

CTT TAC TTT TTG GGT AAC CTG AAC CGC CTC ATC AAG GCC AAA TGG TTG   1646
Leu Tyr Phe Leu Gly Asn Leu Asn Arg Leu Ile Lys Ala Lys Trp Leu
            510             515             520

TGG GAT CCC AAC GAG ATC TTC ACA AAC AAA CAG AGC ATC CCT ACT AAA   1694
Trp Asp Pro Asn Glu Ile Phe Thr Asn Lys Gln Ser Ile Pro Thr Lys
            525             530             535

CCT CTT AAG GAG CCC AAG CAG ACG AAA TAGTAGGTCA CAATTAGTCA         1741
Pro Leu Lys Glu Pro Lys Gln Thr Lys
            540             545

TCGACTGAAG TGCAGCACTT GTCGGATACG GCGTGATGGT TGCTTTTTAT AAACTTGGTA 1801
```

(2) INFORMATION FOR SEQ ID NO: 31:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 546 amino acids
            (B) TYPE: amino acid
            (D) TOPOLOGY: linear

       (ii) MOLECULE TYPE: protein

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

```
Met Ala Thr Leu Pro Gln Lys Asp Pro Gly Tyr Ile Val Ile Asp Val
1               5               10              15

Asn Ala Gly Thr Ala Asp Lys Pro Asp Pro Arg Leu Pro Ser Met Lys
            20              25              30

Gln Gly Phe Asn Arg Arg Trp Ile Gly Thr Asn Ile Asp Phe Val Tyr
        35              40              45

Val Val Tyr Thr Pro Gln Gly Ala Cys Thr Ala Leu Asp Arg Ala Met
    50              55              60

Glu Lys Cys Ser Pro Gly Thr Val Arg Ile Val Ser Gly Gly His Cys
 65             70              75              80

Tyr Glu Asp Phe Val Phe Asp Glu Cys Val Lys Ala Ile Ile Asn Val
                85              90              95
```

Thr Gly Leu Val Glu Ser Gly Tyr Asp Asp Asp Arg Gly Tyr Phe Val
            100   105   110

Ser Ser Gly Asp Thr Asn Trp Gly Ser Phe Lys Thr Leu Phe Arg Asp
            115   120   125

His Gly Arg Val Leu Pro Gly Gly Ser Cys Tyr Ser Val Gly Leu Gly
    130   135   140

Gly His Ile Val Gly Gly Gly Asp Gly Ile Leu Ala Arg Leu His Gly
145   150   155   160

Leu Pro Val Asp Trp Leu Ser Gly Val Glu Val Val Val Lys Pro Val
                165   170   175

Leu Thr Glu Asp Ser Val Leu Lys Tyr Val His Lys Asp Ser Glu Gly
            180   185   190

Asn Asp Gly Glu Leu Phe Trp Ala His Thr Gly Gly Gly Gly Gly Asn
            195   200   205

Phe Gly Ile Ile Thr Lys Tyr Tyr Phe Lys Asp Leu Pro Met Ser Pro
210   215   220

Arg Gly Val Ile Ala Ser Asn Leu His Phe Ser Trp Asp Gly Phe Thr
225   230   235   240

Arg Asp Ala Leu Gln Asp Leu Leu Thr Lys Tyr Phe Lys Leu Ala Arg
                245   250   255

Cys Asp Trp Lys Asn Thr Val Gly Lys Phe Gln Ile Phe His Gln Ala
            260   265   270

Ala Glu Glu Phe Val Met Tyr Leu Tyr Thr Ser Tyr Ser Asn Asp Ala
            275   280   -   285

Glu Arg Glu Val Ala Gln Asp Arg His Tyr His Leu Glu Ala Asp Ile
290   295   300

Glu Gln Ile Tyr Lys Thr Cys Glu Pro Thr Lys Ala Leu Gly Gly His
305   310   315   320

Ala Gly Trp Ala Pro Phe Pro Val Arg Pro Arg Lys Arg His Thr Ser
                325   330   335

Lys Thr Ser Tyr Met His Asp Glu Thr Met Asp Tyr Pro Phe Tyr Ala
            340   345   350

Leu Thr Glu Thr Ile Asn Gly Ser Gly Pro Asn Gln Arg Gly Lys Tyr
            355   360   365

Lys Ser Ala Tyr Met Ile Lys Asp Phe Pro Asp Phe Gln Ile Asp Val
370   375   380

Ile Trp Lys Tyr Leu Thr Glu Val Pro Asp Gly Leu Thr Ser Ala Glu
385   390   395   400

Met Lys Asp Ala Leu Leu Gln Val Asp Met Phe Gly Gly Glu Ile His
                405   410   415

Lys Val Val Trp Asp Ala Thr Ala Val Ala Gln Arg Glu Tyr Ile Ile
            420   425   430

Lys Leu Gln Tyr Gln Thr Tyr Trp Gln Glu Glu Asp Lys Asp Ala Val
        435   440   445

Asn Leu Lys Trp Ile Arg Asp Phe Tyr Glu Glu Met Tyr Glu Pro Tyr
    450   455   460

Gly Gly Val Pro Asp Pro Asn Thr Gln Val Glu Ser Gly Lys Gly Val
465   470   475   480
Phe Glu Gly Cys Tyr Phe Asn Tyr Pro Asp Val Asp Leu Asn Asn Trp
            485   490   495

Lys Asn Gly Lys Tyr Gly Ala Leu Glu Leu Tyr Phe Leu Gly Asn Leu
            500   505   510

Asn Arg Leu Ile Lys Ala Lys Trp Leu Trp Asp Pro Asn Glu Ile Phe
        515   520   525

Thr Asn Lys Gln Ser Ile Pro Thr Lys Pro Leu Lys Glu Pro Lys Gln
530   535   540

Thr Lys
545

(2) INFORMATION FOR SEQ ID NO: 32:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: other nucleic acid

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

ACCAAGTTTA TAAAAAGCAA CCATCAC

27

(2) INFORMATION FOR SEQ ID NO: 33:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: other nucleic acid

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

ATGAATTCGT GGGTCGAAGA GCCC   24

(2) INFORMATION FOR SEQ ID NO: 34:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 33 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: other nucleic acid

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

CAGGAATTCA TATGGCTACT CTTCCCCAGA AAG   33